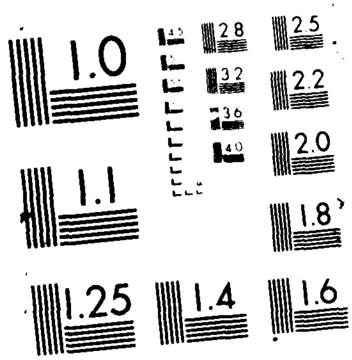
AD-A185 907 EXPLICIT CONSTRUCTION AND ANALYSIS OF SOME EFFICIENT 1/1
EMPIRICAL PRODUCTION..(U) TEXAS UNIV AT AUSTIN CENTER
FOR CYBERNETIC STUDIES A CHARNES ET AL. JUN 87 CCS-582
UNCLASSIFIED N00014-86-C-0398 F/G 12/4 NL

END


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

A185 907

(12)

Research Report CCS 582

EXPLICIT CONSTRUCTION AND ANALYSIS OF SOME EFFICIENT EMPIRICAL PRODUCTION FUNCTIONS IN DATA ENVELOPMENT ANALYSIS

by
A. Charnes
W. W. Cooper
D. B. Learner*
F.Y. Phillips*
J. Rousseau*

June 1987

*MRCA Information Services

This research was partly supported by National Science Foundation Grants SES-8408134 and SES-5520806 and by Office of Naval Research contracts N00014-86-C-0398 and N00014-82-K-0389 with the Center for Cybernetic Studies, The University of Texas at Austin and partly supported by the IC[2] Institute of The University of Texas at Austin. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director

College of Business Administration, 5.202
The University of Texas at Austin
Austin, Texas 78712-1177
(512) 471-1821

ABSTRACT

In contrast to classical econometric work which only tests data for consistency with a special class of production functions, new theory and explicit construction by Data Envelopment Analysis of the empirical Pareto-Koopmans efficient production function is developed for data sets which satisfy two conditions met in all previous real applications of DEA known to the authors. The construction requires no additional computation beyond that of the DEA tests.

KEY WORDS

Data Envelopment Analysis

Pareto-Koopmans efficiency

Multi-criteria Programming

Empirical Efficient Production Functions

1. Introduction

The "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions" by Charnes, Cooper, Golany, Seiford and Stutz (1985), hereinafter referred to as "Foundations," initiated a basic theoretical analysis and interpretation of Data Envelopment Analysis (DEA) as an approach for analyzing and evaluating the (relative) efficiency of performance of DMUs (Decision Making Units) by reference to production possibility sets specified directly instead of implicitly through point-to-set mappings required to satisfy properties imposed by a priori axioms as in Shephard (1970). Additionally, the DEA approach and orientation involved production possibility sets specified or estimated directly from empirical data which conform to real managerial possibilities and from which the efficient empirical production function is to be specified insofar as it can be known from the data and possibility set assumed in contrast to the emphasis and approaches associated with the abstract axiomatic econometric tradition.

In section 2, we bring forth the differences in key preoccupations of research in economic production theory with those of DEA by references to papers of Hanoch and Rothschild (1972), of Diewert and Parkan (1983) and, in still another direction by Färe, Grosskopf and Lovell (1985) and Debreu (1951).

In section 3 we develop basic concepts of and theorems for geometric elucidation and analysis of the empirical efficient production functions of DEA including a new lemma on optimal solutions to the general linear programming problem (for arbitrary ordered fields of scalars). Then, restricting ourselves to two conditions on empirical data which have held in every DEA application we have made, we show that the mathematical structure of the efficient empirical function is so simplified that it is available immediately in analytic form without further computational work on top of the

DEA tests. The efficient empirical production is then, moreover, piece-wise linear and continuous from piece to piece.

2. DEA and Classic Economic Approaches

While the orientation in DEA is toward construction of efficient production functions with best possible conformance to observed empirical data, approaches taken and objectives in classical approaches are quite different. This may be made more concrete by reference to a very early and a very recent article which are directed toward testing whether observed data conform to requirements specified in formal economic theories of production e.g. of the Shephard type. Thus, Hanoch and Rothschild (1972) is an early example of the latter approach to determine whether empirical observations have the desired properties via linear programming techniques. Diewert and Parkan (1983) provide a recent example in this same tradition. In both cases, the orientation is toward applying linear programming formulations to different bodies of data in order to ascertain whether they are globally consistent with the properties postulated for the production functions that are used in economic theory.

As noted in Diewert and Parkan, p. 131, for example, "it is important to know whether a failure to satisfy economic regularity conditions [in a statistically fitted function] is due to the use of an inappropriate functional form or whether the given data are simply inconsistent with any functional form satisfying the appropriate regularity conditions." In fact, in the latter vein, Hanoch and Rothschild, p. 273, report that a body of data used by others for purposes of statistical estimation is dangerous because their tests show that these data fail to satisfy the regularity conditions required by the economic theory of production.

In contrast, DEA as described in Foundations 1985 starts from observations on empirical data and proceeds to determine locally, with each DMU, a "facet" spanned by efficient DMUs from which a piece of the efficient empirical function (of a much wider class than those in this economic literature) together with output shortfalls and input surpluses is determined rather than merely a single efficiency score. The DEA approach secures a Pareto-Koopmans efficient function from the very nature of the DEA models, which, as shown in Foundations, represent the Charnes-Cooper test for Pareto-Koopmans optimality (efficiency).

We can illuminate other differences between DEA and these other approaches by distinguishing between "pure predictions" and "control predictions." As is noted on pp. 101 ff. in Charnes, Cooper, Phillips and Learner (1985), the usual least squares-regression approach is pointed toward pure prediction in the sense of its ability to extrapolate or interpolate the behavior of the dependent variable (i.e., the regressant) without attempting to alter underlying patterns of behavior exhibited by past data. DEA, on the other hand, provides formulas for projecting inefficient DMUs onto the efficiency frontier as in Charnes, Cooper and Rhodes (1978). The presumption is that the observed behavior for these adjusted DMUs can be altered to conform with these projections. In contrast to this "control" approach, the analyses development in Diewert and Parkan (1983) is restricted to measures of efficiency (efficiency scores) for passive use and, in addition, the failure to allow for the presence of non-zero slack renders these measures unsuitable for use in effecting such projections.

Modifications of the CCR production possibility sets, hence, the DEA model employed, may be needed of course, when some of the indicated projections cannot be attained because some components in the input vectors are partially or wholly non-discretionary. Charnes, Cooper, Rousseau and Semple (1987) provide the necessary modifications to the model to conform to managerially possible production for both cases. A treatment had been suggested in Banker and Morey (1986) for providing modified efficiency scores and thereby efficient DMUS when inputs and outputs can be identified into two classes according to whether their values are (or are not) completely fixed exogenously. It has not been shown, however, that their modified test correctly yields Pareto-Koopmans efficiency or that a correct production possibility set is effectively achieved. No such developments are to be found in the traditions of the literature covered by Hanoch and Rothschild and Diewert and Parkan which may be thought of as being more oriented toward "understanding" rather than the "understanding for use" point of view described in Charnes, Cooper, Learner and Phillips (1985).

The tradition represented by Hanoch and Rothschild and Diewert and Parkan is not the only one that can be identified in the literature of econometrics. Starting with the classic article by Farrell (1957)[1], another related series of efforts in econometrics is directed toward developing scalar overall as well as component measures of technical, scale and allocative efficiencies. Recent articles in this tradition may be found in Färe and Lovell (1978) as well as in Russell (1985) with a comprehensive and detailed treatment being available in Färe, Grosskopf and Lovell (1985).

Starting with the "CCR ratio form" as given in Charnes, Cooper and Rhodes (1978) the DEA literature has emphasized explicitly developed methods for locating sources and estimating amounts of inefficiency directly from the data. This has continued into subsequent formulations which embody the DEA principle in different forms (and different production possibility sets) such as the "invariant multiplicative form" of Charnes, Cooper, Seiford and Stutz (1983) and the "additive model" of Charnes, Cooper, Golany, Seiford and Stutz (1985) as well as the "extended additive model" as

---

[1] See also the earlier article by Debreu (1951).

given in Charnes, Cooper, Rousseau and Semple (1987). In the following we restrict ourselves to the first three, the necessary elaborations for the latter being in progress.

3 DEA Geometry and Efficient Function Construction

The CCR ratio model may be rendered as the dual (non-Archimedean) linear programs:

$$
\text{(CCR)}\qquad
\begin{array}{ll}
\text{(R)} & \\
\max h = \mu^T y_0 & \\
\qquad \nu^T x_0 = 1 & \\
\qquad \mu^T Y - \nu^T X \leq 0 & \\
\qquad -\mu^T \leq -\epsilon e^T & \\
\qquad -\nu^T \leq -\epsilon e^T &
\end{array}
\qquad\qquad
\begin{array}{l}
\text{(DEA)} \\
\min \theta - \epsilon e^T s^+ - \epsilon e^T s^- \\
Y\lambda - s^+ = y_0 \\
\theta x_0 - X\lambda - s^- = 0 \\
\\
\lambda, s^+, s^- \geq 0
\end{array}
$$

where $Y \triangleq [y_1, \ldots, y_n]$, $X \triangleq [x_1, \ldots, x_n]$ are matrices of positive vectors and ϵ is a non-Archimedean infinitesimal.

The additive model is

$$
\text{(A)}\qquad
\begin{array}{lll}
\min & -e^T s^+ - e^T s^- & \\
& Y\lambda - s^+ & = y_0 \\
& -X\lambda \qquad - s^- & = -x_0 \\
& e^T\lambda & = 1 \\
& \lambda, s^+, s^- \geq 0 &
\end{array}
$$

If one wishes this value to be independent of the units of measurement we alter the functional to

$$
\text{(A.2)}\qquad e^T D^{-1}(y_0)s^+ + e^T D^{-1}(x_0)s^- = d^w(s^+, s^-)
$$

where $D(y_0)$, $D(x_0)$ are diagonal matrices with the y_0 or x_0 component entries. (If some components of y_0 or x_0 are zero we use the unique Moore-Penrose generalized inv ~se which has zeros instead of reciprocals for the zero components.) For either functional we obtain the same DMUS as efficient since

$$\text{(A.2.1)} \qquad \alpha(e^T s^+ + e^T s^-) \leq e^T D^{-1}(y_0) s^+ + e^T D^{-1}(x_0) s^- \leq \beta(e^T s^+ + e^T s^-)$$

where α, β are respectively the minimum and maximum of the non-zero entries of both $D^{-1}(y_0)$ and $D^{-1}(x_0)$ and therefore $\alpha, \beta > 0$. It is useful to have the (A.1) functional form both for theory and practice since the efficiency hyperplane it defines $(e^T s^+ + e^T s^- = 0)$ does not depend on any particular x_0, y_0.

The Invariant Multiplicative model is formally the same as the Additive model but with input and output components replaced by their logarithms.

As may be noted (cf. Foundations), the DEA side for each corresponds to a Charnes-Cooper test for Pareto-Koopmans efficiency over the relevant production possibility set. Formally, the test could be applied to arbitrary combinations of input-output vector pairs which have no correspondence to actual production function possibilities let alone efficient ones. Thus in all our DEA applications care has been taken to choose inputs and outputs with the expectation that for each single output increases in inputs should not cause decrease in the output. As shown in Foundations, this is a property of Pareto-Koopmans efficient empirical functions with a single output and generalizations to multiple output situations are also given there. The efficient multiple output and input functions need not be isotone, although for any pair of observed inputs there is a cone of directions in output space for which isotonicity holds in these ouput directions.

The DEA tests associate with each DMU tested efficient DMUs whose convex (conical for the CCR model) combinations of input-output vectors form a "facet" of efficient input-output vectors. Each facet corresponds to a piece of the efficient empirical function.

To generate the efficient DMUs associated with a particular DMU we need first the following new lemma on optimal solutions to the general linear programming problem (with an arbitrary ordered field of scalars)

$$\begin{aligned} \max \quad & c^T\lambda \\ & P\lambda = P_0 \\ & \lambda \geq 0 \end{aligned}$$

This problem may be rewritten in terms of a basic solution $\bar{\lambda}^T = (\bar{\lambda}_B{}^T, 0)$ and associated decompositions $c^T = (c_B{}^T, c_N{}^T)$, $P = [B, N]$, $\lambda^T = (\lambda_B{}^T, \lambda_N{}^T)$ as

$$\begin{aligned} \max \quad & c_B{}^T\lambda_B + c_N{}^T\lambda_N \\ & B\lambda_B + N\lambda_N = P_0 \\ & \lambda_B, \lambda_N \geq 0 \end{aligned}$$

The lemma may now be stated:

<u>Lemma</u>: If $\lambda_j > 0$ in some optimal solution (basic or not)

then in every <u>optimal</u> basic solution tableau its reduced cost (shadow price) is zero.

<u>Proof</u>: For every basis B, the constraint equation can be solved to yield

$$\lambda_B = B^{\#}P_0 - B^{\#}N\lambda_N$$

where $B^{\#}$ is any left inverse of B. Thus

$$c_B{}^T\lambda_B + c_N{}^T\lambda_N = c_B{}^TB^{\#}P_0 - (c_B{}^TB^{\#}N - c_N{}^T)\,\lambda_N$$

If B now corresponds to an optimal basic solution $\bar{\lambda}^T = (\bar{\lambda}_B{}^T, 0)$, we have

$$c^T\bar{\lambda}^T = c_B{}^T\bar{\lambda}_B = c_B{}^T B^{\#} P_0$$

as the optimal value.

Thus

$$c_B{}^T B^{\#} P_0 = c^T B^{\#} P_0 - (c_B{}^T B^{\#} N - c_N{}^T)\, \lambda_N$$

for any optimal solution $\lambda^T = (\lambda_B{}^T, \lambda_N{}^T)$ and thereby

$$(c_B{}^T B^{\#} N - c_N{}^T)\, \lambda_N = 0$$

But in an optimal basic solution tableau, the reduced cost vector $(c_B{}^T B^{\#} N - c_N{}^T) \geq 0$

Since also $\lambda_N \geq 0$, we have a sum of non-negatives equal to zero in this equation and thereby

$$(c_B{}^T B^{\#} N_j - c_j{}^T)\, \lambda_j = 0, \qquad \text{for each } j \in N$$

Thus $\lambda_j > 0$ implies $c_B{}^T B^{\#} N_j - c_j{}^T = 0.$

Q.E.D

As an illustration consider

$$\max \quad x_3 \text{ subject to } 0 \leq x_i \leq 1, \qquad i = 1, 2, 3$$

Introducing slacks to convert the inequalities to equations, an optimal tableau is

	$\vec{c}$	0	0	0	1	0	0	0
c_B	B	P_0	x_1	x_2	x_3	s_1	s_2	s_3
0	s_1	1	1			1		
0	s_2	1		1			1	
1	x_3	1			1			1
	$\vec{\zeta}$	1	0	0	0	0	0	1

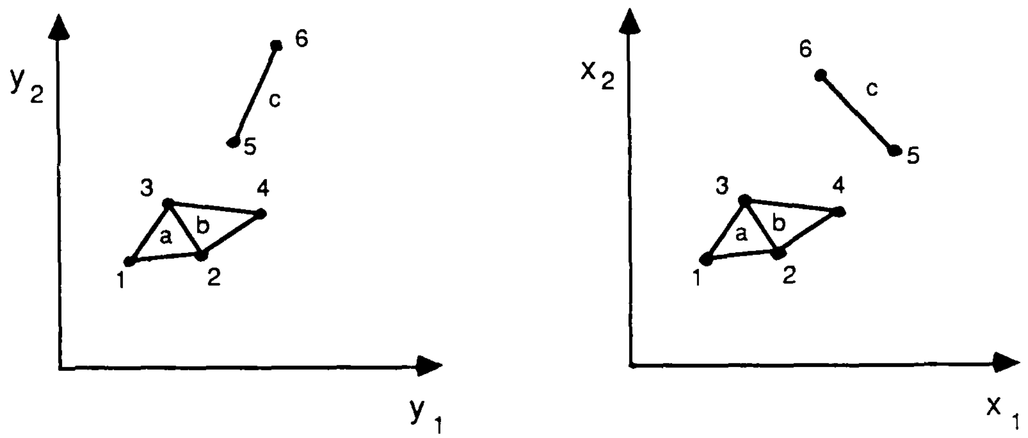
Note that x_1, x_2, x_3, s_1, s_2, may all occur in optimal solutions whereas s_3 cannot e.g. (0,0,1,1,1,0), (1,0,1,0,1,0), (0,1,1,1,0,0), (1,1,1,0,0,0). Further note (1,1,1,0,0,0) is <u>not</u> an adjacent extreme point to the optimal basic solution whose tableau we have used.

Since DMUs whose reduced cost is zero in an optimal basic solution (to the C^2 test) correspond to alternate basic optimal solutions, they are efficient as well as those in the basic optimal solution (cf. Foundations). Thus we have

<u>Theorem 1</u>: The DMUs appearing in an optimal basic solution to a DEA test together with those whose reduced costs are zero are the efficient generators of the facet corresponding to the DMU being evaluated.

To illustrate empirical efficient function development, consider the following 2 input, 2 output graphs for the additive model:

<u>Figure 1</u>



Here 6 efficient DMUs input-output points are plotted in separate input and output spaces rather than the input-output product space of the C^2 test. We suppose they are the generators of the facets a, b, c corresponding to 3 inefficient DMUs not shown. Note that so far as the empirical efficient production function is concerned we have it defined only over the input domain labeled by a, b, c, a geometrical situation which often occurs with empirical distribution functions in statistics.

Now we utilize some of our experiences in actual uses of DEA to remark that in every application we have encountered since the beginning of DEA, the following properties were present. Every facet's generators has consisted of DMUs whose

(a) input vectors are linearly independent

(b) number does not exceed the number of inputs

Theorem 2: If the DMU observation set has properties (a) and (b) then

(i) the efficient empirical production function is uniquely defined over input simplexes corresponding to each facet, i.e., for any input point which is a convex combination of the facet generators its output point is the same convex combination of the generator outputs,

(ii) it is linear on these facets

and,

(iii) continuous when crossing the boundaries between two facets.

Proof: Because of properties (a) and (b) the convex hull of the facet generators in the product input-output space defines a set of input-output points which are efficient and for which any input point in the convex hull of the inputs has a unique representation in input space in terms of the input vectors of the facet generators since by (a) and (b) this convex hull is a simplex in input space. The value of the function in the output space at this input point is the output vector which is the same convex combination of the output vectors of the facet generators. Thus this function is linear over this input vector simplex.

The boundary between two adjacent facets consists of convex combinations of the subset of input-output points which are in both facets. The value of the production function at a point on the boundary is thus obtained from the same (unique) combination of the DMUs generating these boundaries and hence its value is the same for both adjacent facets. Therefore, the empirical production function is continuous across the boundary between adjacent facets.

Q.E.D.

Thus each linear piece of the function is easily specified analytically without further computation and a path is open for further extensions (and uses) of DEA. These extensions can move from considerations of efficiency to considerations of effectiveness and studies of stability or senstivity can be addressed with relative ease by the mathematical programming methods that are available for these purposes. See, e.g., Charnes, Cooper, Lewin, Morey and Rousseau (1985). This work together with that of Charnes and Neralic (1986) also bears on the problems of enlargement of the empirical Pareto-Koopmans efficient production function to input vector domains which, for example, may fill in gaps between input domain pieces as shown in Figure 1. Such problems were brought forth and discussed in Foundations. Further contributions to its resolution would further illuminate possible relations between the Hanoch, Rothschild, Diewert and Parkan results and DEA contributions.

REFERENCES

Banker, R. and R. Morey, "Efficiency Analysis with Exogenously Fixed Inputs and Outputs," Operations Research (1986), pp. 513-521.

Charnes, A., W.W. Cooper, B. Golany, L. Seiford and J. Stutz (1985), "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions," Journal of Econometrics 30, pp. 91-107.

Charnes, A., W.W. Cooper, D.B. Learner and F.Y. Phillips (1985), "Management Science and Marketing Management," Journal of Marketing, pp. 93-105.

Charnes, A., W.W. Cooper, A.Y. Lewin, R.C. Morey and J. Rousseau (1985), "Sensitivity and Stability Analysis in DEA," Annals of Operations Research, pp. 139-150.

Charnes, A., W.W. Cooper and E. Rhodes, "Measuring the Efficiency of Decision Making Units," (1978), European Journal of Operational Research, pp. 429-444.

Charnes, A., W.W. Cooper, J. Rousseau and J. Semple (submitted, 1987), "Data Envelopment Analysis and Axiomatic Notions of Efficiency and Reference Sets," Management Science.

Charnes, A., W.W. Cooper, L. Seiford and J. Stutz (1983), "Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopments," (1983), Operations Research Letters, pp. 101-103.

Charnes, A., W.W. Cooper, L. Seiford and J. Stutz(1983), "A Structure for Characterizing and Classifying Efficiencies and Inefficiencies in Data Envelopment Analysis," CCS Research Report 512 (Austin: The University of Texas Center for Cybernetic Studies).

Charnes, A., W.W. Cooper, and R.M. Thrall, (1986), "Identifying and Classifying Scale and Technical Efficiencies and Inefficiencies in Observed Data Via Data Envelopment Analysis," Operations Research Letters, pp. 105-110.

Charnes, A., and L. Neralic, (1986), "Sensitivity Analysis in DEA," CCS Research Report 530, "Sensitivity Analysis in DEA - Part II," CCS Research Report 536, "Sensitivity in DEA - Part III," CCS Research Report 542, "Sensitivity in DEA - Part IV," CCS Research Report 543, "Sensitivity in DEA - Part V," CCS Research Report 544, "Sensitivity in DEA - Part VI," CCS Research Report 545, (Austin: The University of Texas Center for Cybernetic Studies).

Debreu, G., "The Coefficient of Resource Utilization," Econometrica (1951), pp 273-292.

Diewert, W.E. and C. Parkan, "Linear Programming Tests of Regularity Conditions for Production Functions," (1983), R. Henn, K. Neumann, and R.W. Shephard, eds., Quantitative Studies on Production and Prices (Wurtzburg: Physica-Verlag), pp. 131-158.

Färe, R. and C.A.K. Lovell, "Measuring the Technical Efficiency of Production," (1978), Journal of Economic Theory, pp. 150-162.

Farrell, M.J., "The Measurement of Productive Efficiency," (1957), Journal of the Royal Statistical Society, Series A, pp. 251-281.

Hanoch, G. and M. Rothschild, "Testing the Assumptions of Production Theory: A Nonparametric Approach," (1972), Journal of Political Economy, pp. 256-275.

Russell, R.R., "Measures of Technical Efficiency," (1985), Journal of Economic Theory, pp. 109-126.

Russell, R.R., "On the Axiomatic Approach to the Measurement of Technical Efficiency," (1987), Working Paper, University of California at Riverside 92521, pp. 1-22.

Shephard, R., (1970), Theory of Cost and Production Functions (Princeton, NJ: Princeton University Press).

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CCS 582	2. GOVT ACCESSION NO. ADA186023	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EXPLICIT CONSTRUCTION AND ANALYSIS OF SOME EFFICIENT EMPIRICAL PRODUCTION FUNCTIONS IN DATA ENVELOPMENT ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER CCS 582
7. AUTHOR(s) A. Charnes W.W. Cooper D.B. Learner, MRCA Information Services F.Y. Phillips, MRCA J. Rousseau, MRCA		8. CONTRACT OR GRANT NUMBER(s) N00014-86-C-0398 N00014-82-K-0389
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Cybernetic Studies, UT Austin Austin, Texas 78712		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 434) Washington, D.C.		12. REPORT DATE June 1987
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Data Envelopment Analysis, Pareto-Koopmans efficiency, Multi-criteria Programming, Empirical Efficient Production Functions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In contrast to classical econometric work which only tests data for consistency with a special class of production functions, new theory and explicit construction by Data Envelopment Analysis of the empirical Pareto-Koopmans efficient production function is developed for data sets which satsify two conditions met in all previous real applications of DEA known to the authors. The construction requires no additional computation beyond that of the DEA tests.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

12-87

DTIC